

experiment7

May 2, 2025

```
[2]: import pandas as pd
```

```
[3]: import matplotlib.pyplot as plt
```

```
[4]: df=pd.read_csv(r"C:\Users\advik\OneDrive\Desktop\experiment4.csv a.csv")
```

```
[5]: df
```

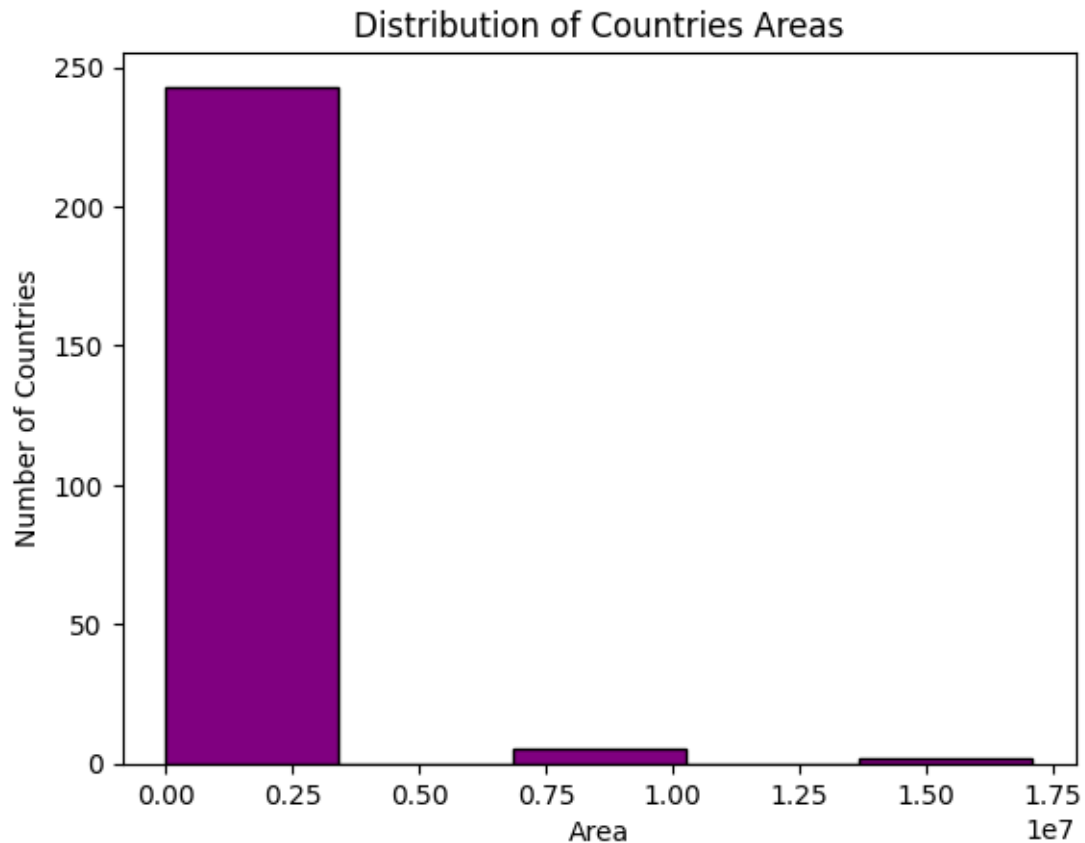
```
[5]:
```

	name	capital	population	area	region
0	South Georgia	King Edward Point	30	3903.0	Antarctic
1	Grenada	St. George's	112519	344.0	Americas
2	Switzerland	Bern	8654622	41284.0	Europe
3	Sierra Leone	Freetown	7976985	71740.0	Africa
4	Hungary	Budapest	9749763	93028.0	Europe
..
245	Belgium	Brussels	11555997	30528.0	Europe
246	Israel	Jerusalem	9216900	20770.0	Asia
247	New Zealand	Wellington	5084300	270467.0	Oceania
248	Nicaragua	Managua	6624554	130373.0	Americas
249	Anguilla	The Valley	13452	91.0	Americas

[250 rows x 5 columns]

1 Histogram plot

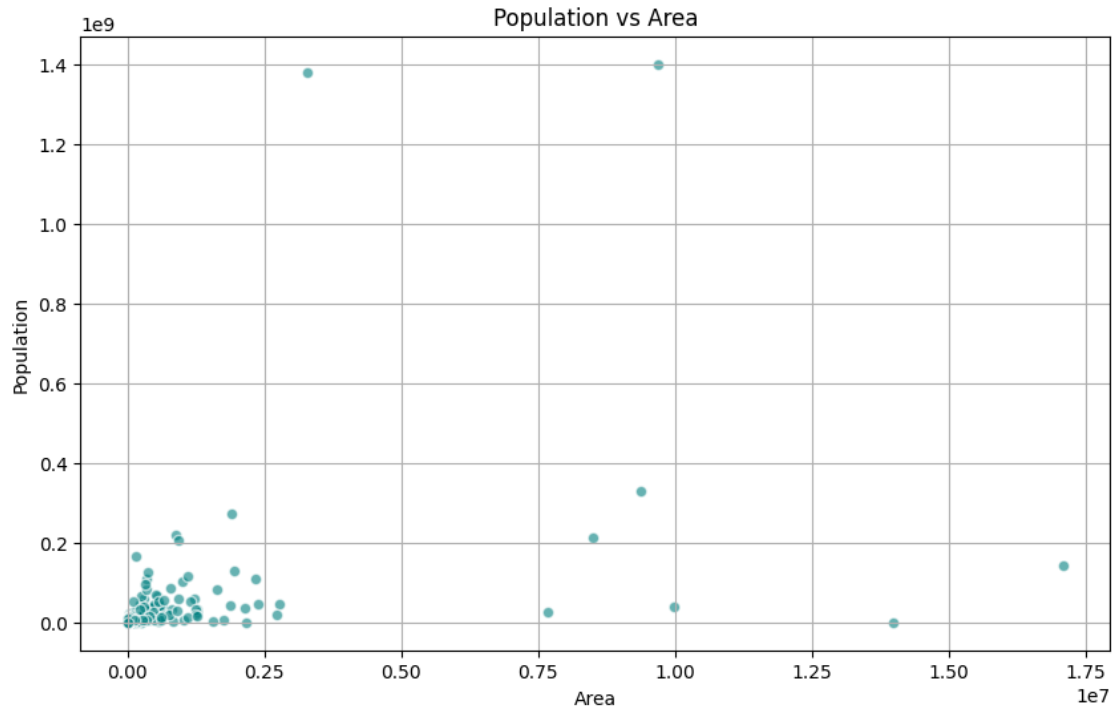
```
[6]: plt.hist(df['area'], bins=5, color='purple', edgecolor='black')
plt.xlabel('Area')
plt.ylabel('Number of Countries')
plt.title('Distribution of Countries Areas')
plt.show()
```



INSIGHTS: Most countries have small land areas: Over 250 countries fall into the smallest area range (leftmost bar). Very few large countries: As area increases, the number of countries drastically decreases. Only a few countries fall into the larger area categories. The vast majority of the world's countries are relatively small in land area, while only a few occupy very large territories.

2 Scatter plot

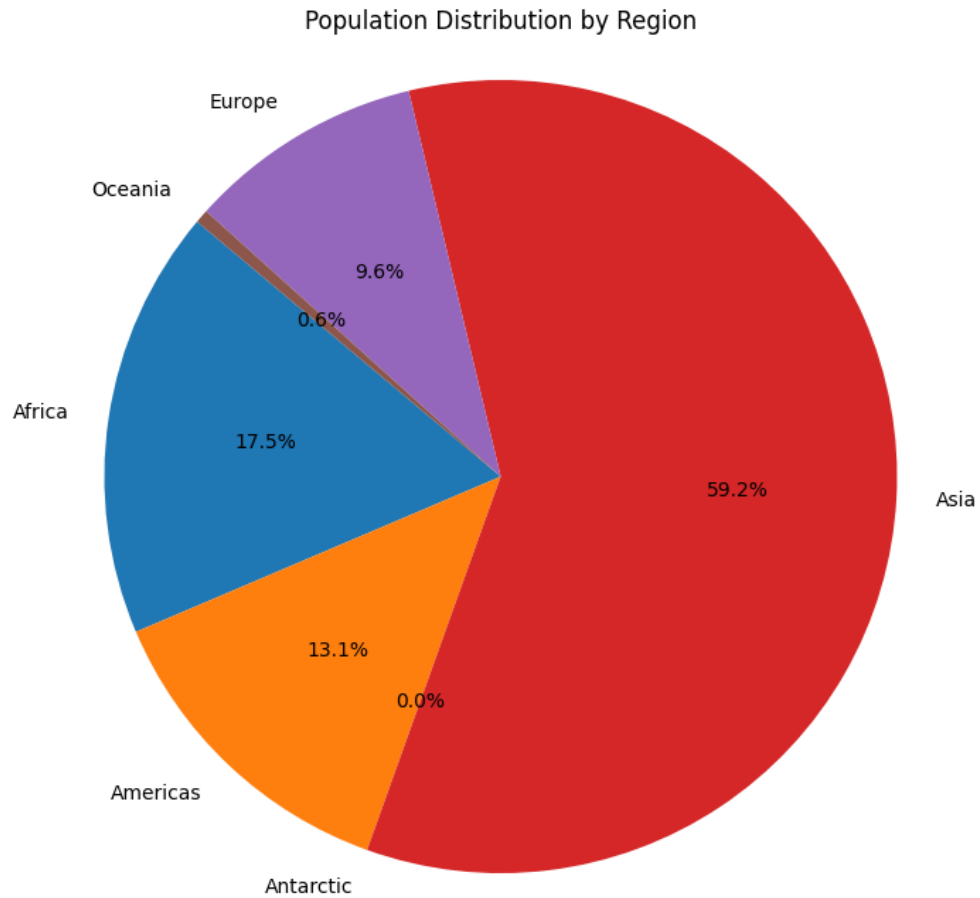
```
[ ]: plt.figure(figsize=(10,6))
plt.scatter(df['area'],
df['population'], alpha=0.6, edgecolors='w', color='teal'),
plt.title('Population vs Area')
plt.xlabel('Area')
plt.ylabel('Population')
plt.grid(True)
plt.show()
```



INSIGHTS: This scatter plot titled “Population vs Area” shows the relationship between a country’s land area and its population. There’s no strong correlation between area and population. Some countries have very large populations but small areas, and vice versa. A few countries (likely China and India) have both very large populations (over 1.4 billion).

3 Pie plot

```
[ ]: region_population = df.groupby('region')['population'].sum()
plt.figure(figsize=(10, 8))
plt.pie(region_population, labels=region_population.index, autopct='%1.1f%%',
        ↪startangle=140)
plt.title('Population Distribution by Region')
plt.axis('equal')
plt.show()
```



INSIGHTS: Asia has the largest population share (59.2%), followed by Africa (17.5%) and the Americas (13.1%). Europe holds 9.6%, while Oceania and Antarctica have minimal or no population.